

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 1, 2005, 14:19:08 ; Search time 200 Seconds
(without alignments)
696.203 Million cell updates/sec

Title: US-10-629-329A-2
Perfect score: 1322
Sequence: 1 MSGCDAGGDCSRRCGAQD.....SMKKVGLDPSQLPVGENGIV 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1314	99.4	242	Q9Y318	Q9Y318 homo sapien
2	1307	98.9	242	Q96GX9	Q96GX9 homo sapien
3	1296	98.0	242	Q8WVU2	Q8WVU2 homo sapien
4	1295	98.0	242	Q96HK2	Q96HK2 homo sapien
5	1239.5	93.8	241	Q9WVQ5	Q9WVQ5 mus musculus
6	1229.5	93.0	241	Q8BP46	Q8BP46 mus musculus
7	1040	78.7	239	Q6NU29	Q6NU29 xenopus lae
8	1040	78.7	239	AAH68773	AAH68773 xenopus lae
9	1005	76.0	204	Q6PJX6	Q6PJX6 homo sapien
10	1005	76.0	204	AAH10133	AAH10133 homo sapi
11	830	62.8	153	Q9H528	Q9H528 homo sapien
12	790	59.8	227	Q9VY93	Q9VY93 drosophila
13	671	50.8	277	Q7PS09	Q7PS09 anopheles g
14	617	46.7	220	Q7PG25	Q7PG25 anopheles g
15	608.5	46.0	238	Q6CBB0	Q6CBB0 yarrowia li
16	559	42.3	507	Q9FN41	Q9FN41 arabidopsis
17	559	42.3	507	AAU06425	AAU06425 arabidops
18	549	41.5	258	Q7SF46	Q7SF46 neurospora
19	549	41.5	258	CAE76315	CAE76315 neurospor
20	542.5	41.0	265	Q6BIX5	Q6BIX5 debaryomyce
21	495	37.4	244	YJ24_YEAST	YJ24_YEAST
22	492.5	37.3	242	Q75CP5	Q75CP5 ashbya goss
23	492.5	37.3	242	AA511102	AA511102 ashbya go
24	477	36.1	208	Q6FJA5	Q6FJA5 candida gla
25	431.5	32.6	284	Q23261	Q23261 caenorhabdi
26	430	32.5	205	Q6CM29	Q6CM29 kluyveromyc
27	409.5	31.0	221	Q9HE08	Q9HE08 schizosacch
28	289	21.9	240	Q8TA31	Q8TA31 heterodera
29	274.5	20.8	192	Q9UT22	Q9UT22 schizosacch
30	215.5	16.3	209	MTNB_BACSU	Q31668 bacillus su
31	214	16.2	202	Q828L0	Q828L0 streptomyce

32	214	16.2	207	2	Q7V8Y6	Q7V8Y6 prochloroco
33	208	15.7	205	2	Q6DI1G1	Q6DI1G1 erwinia car
34	205.5	15.5	211	2	Q7U4V0	Q7U4V0 synechococc
35	196	14.8	212	2	Q819B6	Q819B6 bacillus ce
36	195	14.8	212	2	Q6HEC7	Q6HEC7 bacillus th
37	195	14.8	212	2	Q731R0	Q731R0 bacillus ce
38	195	14.8	212	2	Q81MJ0	Q81MJ0 bacillus an
39	195	14.8	212	2	AA543007	AA543007 bacillus
40	195	14.8	212	2	AA733374	AA733374 bacillus
41	193	14.6	204	2	Q884P3	Q884P3 pseudomonas
42	192.5	14.6	208	1	YJ79_AQUAE	YJ79_AQUAE
43	170.5	12.9	249	2	Q8EXC1	Q8EXC1 leptospira
44	166.5	12.6	249	2	Q75FG3	Q75FG3 leptospira
45	166.5	12.6	249	2	AA572252	AA572252 leptospir

ALIGNMENTS

RESULT 1

Q9Y318

ID Q9Y318 PRELIMINARY; PRT; 242 AA.

AC Q9Y318; DT 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-WAR-2002 (Tremblrel. 20, Last annotation update)

DE CGI-29 protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20272150; PubMed=10810093;

RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;

RT "Identification of novel human genes evolutionarily conserved in

RT Caenorhabditis elegans by comparative proteomics."

RL Genome Res. 10:703-713(2000).

DR EMBL; AF132963; AAD27738.1; -.

DR InterPro; IPR001303; Aldolase II_N.

DR Pfam; PF00596; Aldolase II; 1_.

SQ SEQUENCE 242 AA; 27012 MW; 7B99194024C77D5B CRC64;

Query Match 99.4%; Score 1314; DB 2; Length 242;

Best Local Similarity 99.6%; Pred. No. 6.3e-110;

Matches 241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSGCDAGGDCSRRCGAQDKHEPRYLIPELCKQFYHLGWVTGTGGGSLKHGDEITYAP 60

Db 1 MSGCDAGGDCSRRCGAQDKHEPRYLIPELCKQFYHLGWVTGTGGGSLKHGDEITYAP 60

QY 61 SGVQKERIQEDMFYCDINEXDISGSPSKLKKSQCTPLFNAYTRMGAGAVIHTHKA 120

Db 61 SGVQKERIQEDMFYCDINEXDISGSPSKLKKSQCTPLFNAYTRMGAGAVIHTHKA 120

QY 121 AVMTALLPPGREFKITHOEMIKGIKKTSGGYRYDDMLVPIIENITPEEKGLKDRMAHA 180

Db 121 AVMTALLPPGREFKITHOEMIKGIKKTSGGYRYDDMLVPIIENITPEEKGLKDRMAHA 180

QY 181 MNEYPDSCAVLVRHGVVVGWETWEKATKMCCEYDLYFDIAVSMKKVGLDPSQLPVGENG 240

Db 181 MNEYPDSCAVLVRHGVVVGWETWEKATKMCCEYDLYFDIAVSMKKVGLDPSQLPVGENG 240

QY 241 IV 242

Db 241 IV 242

RESULT 2

Q96GX9

ID Q96GX9 PRELIMINARY; PRT; 242 AA.

AC Q96GX9; DT 01-DEC-2001 (Tremblrel. 19, Created)

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CGI-29 protein.
GN Name=MWRP19;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017594; AAH17594.1; -.
DR InterPro; IPR001303; Aldolase II_N.
DR Pfam; PF00596; Aldolase II; 1.
SQ SEQUENCE 242 AA; 27215 MW; F8C4A371E3C63FC2 CRC64;

Query Match 98.0%; Score 1296; DB 2; Length 242;
Best Local Similarity 98.8%; Pred. No. 2.6e-108;
Matches 239; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 MSGCDAGEGDCSRRCGAQDKHPRYLIPELCKQFYHLGWVTGTGGGISLKHGDEIYIAP 60
DB 1 MSGCDAGEGDCSRRCGAQDKHPRYLIPELCKQFYHLGWVTGTGGGISLKHGDEIYIAP 60
QY 61 SGVQKERTIQEDMFVCDINEKDISGPSPSKKLKKSOCTPLFNAYTMRGAGAVIHTSKA 120
DB 61 SGVQKERTIQEDMFVCDINEKDISGPSPSKKLKKSOCTPLFNAYTMRGAGAVIHTSKA 120
QY 121 AVMATLLFPGRFKEITHQEMIKGKCTSGGYRYDDMLVPIENTPEEKLKDRMAHA 180
DB 121 AVMATLLFPGRFKEITHQEMIKGKCTSGGYRYDDMLVPIENTPEEKLKDRMAHA 180
QY 181 MNEYDSCAVLVRHGVYVWGTEWKAKTMCCECYDYLFDAIVSMKKVGLDPSQLPVGENG 240
DB 181 MNEYDSCAVLVRHGVYVWGTEWKAKTMCCECYDYLFDAIVSMKKVGLDPSQLPVGENG 240
QY 241 IV 242
DB 241 IV 242

RESULT 4
Q96HK2 ID Q96HK2 PRELIMINARY; PRT; 242 AA.
AC Q96HK2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CGI-29 protein.
GN Name=MWRP19;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Harte S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009077; AAH09077.1; -.
DR InterPro; IPR001303; Aldolase II_N.
DR Pfam; PF00596; Aldolase II; 1.
SQ SEQUENCE 242 AA; 27125 MW; 9B8D5D1435D6775A CRC64;

Query Match 99.9%; Score 1307; DB 2; Length 242;
Best Local Similarity 99.2%; Pred. No. 2.7e-109;
Matches 240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSGCDAGEGDCSRRCGAQDKHPRYLIPELCKQFYHLGWVTGTGGGISLKHGDEIYIAP 60
DB 1 MSGCDAGEGDCSRRCGAQDKHPRYLIPELCKQFYHLGWVTGTGGGISLKHGDEIYIAP 60
QY 61 SGVQKERTIQEDMFVCDINEKDISGPSPSKKLKKSOCTPLFNAYTMRGAGAVIHTSKA 120
DB 61 SGVQKERTIQEDMFVCDINEKDISGPSPSKKLKKSOCTPLFNAYTMRGAGAVIHTSKA 120
QY 121 AVMATLLFPGRFKEITHQEMIKGKCTSGGYRYDDMLVPIENTPEEKLKDRMAHA 180
DB 121 AVMATLLFPGRFKEITHQEMIKGKCTSGGYRYDDMLVPIENTPEEKLKDRMAHA 180
QY 181 MNEYDSCAVLVRHGVYVWGTEWKAKTMCCECYDYLFDAIVSMKKVGLDPSQLPVGENG 240
DB 181 MNEYDSCAVLVRHGVYVWGTEWKAKTMCCECYDYLFDAIVSMKKVGLDPSQLPVGENG 240
QY 241 IV 242
DB 241 IV 242

RESULT 3
Q8WU2 ID Q8WU2 PRELIMINARY; PRT; 242 AA.
AC Q8WU2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
```

CGI-29 protein.
 DE Name=MRP19; Synonyms=MMR19;
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RP Sha S., Aoki Y., Nishi Y.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M.J., Udell T.B., Toshyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Casavant T.L., Scheetz T.E.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BCO08440; AAH08440.1; -;
 DR EMBL: BC028434; AAH28434.1; -;
 DR MGD: MGI:1926788; Mmrp19.
 DR InterPro: IPR001303; Aldolase_II_N.
 DR Pfam: PF00596; Aldolase_II; 1.
 SQ SEQUENCE 242 AA; 27112 MW; 691C60447B723C1B CRC64;
 Query Match 98.0%; Score 1295; DB 2; Length 242;
 Best Local Similarity 98.3%; Pred. No. 3.2e-108;
 Matches 238; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MSGCDAGEGDCSRRCGAQQKHPHYLIPELCKQFYHLGWVTGTGGGSLKHGDEIYIAP 60
 DB 1 MSGCDAREGDCSRRCGAQQKHPHYLIPELCKQFYHLGWVTGTGGGSLKHGDEIYIAP 60
 QY 61 SGVQKRIQPEDMFVCDINEKDISGSPSKLLKKSQCTPLFMNATMRGAGAVIHTHSA 120
 DB 61 SGVQKRIQPEDMFVCDINEKDISGSPSKLLKKSQCTPLFMNATMRGAGAVIHTHSA 120
 QY 121 AVNATLLFPGRFKEITHQEMIKGIKCTSGGYRYDDMLVPIENTPEEKLKDRMAHA 180
 DB 121 AVNATLLFPGRFKEITHQEMIKGIKCTSGGYRYDDMLVPIENTPEEKLKDRMAHA 180
 QY 181 MNEYPDSCAVLVRHHGVYVWGTEWAKTMCCECYDYLFDIAVSMKKVGLDPSQPVGNG 240
 DB 181 VNEYPDSCAVLVRHHGVYVWGTEWAKTMCCECYDYLFDIAVSMKKVGLDPSQPVGNG 240
 QY 241 IV 242
 DB 241 IV 242
 RESULT 5
 Q9WVQ5 PRELIMINARY; PRT; 241 AA.
 ID Q9WVQ5;
 AC Q9WVQ5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE MMRP19 (CDNA sequence AB028863).

RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,						
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,						
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,						
RA	Krzywinski M.I., Skaleka U., Smalhus D.E., Schnerch A., Schein J.E.,						
RA	Jones S.J., Marra M.A.;						
RT	"Generation and initial analysis of more than 15,000 full-length human						
RT	cDNA sequences.";						
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).						
RN	[3]						
RN	SEQUENCE FROM N.A.						
RC	TISSUE=Embryo;						
RA	Klein S., Strausberg R.;						
RA	Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.						
DW	EMBL; BC068773; AAH68773.1; -.						
KR	Hypothetical protein.						
SQ	SEQUENCE 239 AA; 27172 MW; 690DBAE575AA0CB CRC64;						
Query Match	78.7%; Score 1040; DB 2; Length 239;						
Best Local Similarity	82.0%; Pred. No. 2.8e-85;						
Matches 191; Conservative	19; Mismatches 23; Indels 0; Gaps 0						
QY	10 DCSSRRCGAODKEHPRYLIPELCQFYHLGWVTGTGGISLSKHGDEIYIAPSGVQKERIQ	69					
DB	: :	:					
DB	7 DNCNQTDNAKDKAHPRNLIELCQFYNLGWVTGTGGISLSKYGDEIYIAPSGVQKERIQ	66					
DB	: : : : : :~::~ : : : : : : : : : : : : : : : : :~:: : : : : : : :	:					
QY	70 PEDMFVCIDINEKDISGPSPSKKLKSOCTPLFMNAYTMRGAGAVITHTSKAAMVATLLFP	129					
DB	: :	:					
DB	67 PDDLFFVICIDEKDISCPPPYRNLIKKSQCTPLFMNAYTMRDAGAVITHTSKAAMVATLLMP	126					
DB	: : : : : :~::~ : : : : : : : : : : : : : : : : :~:: : : : : : : :	:					
QY	130 GREFKITHOEMIIGIKKTSGGYRYDDMLVPPIENTPESKGILKDRMAHANVEYPDSCA	189					
DB	: : : : : :~::~ : : : : :~:: : : : :~:: : : : :~:: : : :~:: : : : : :	:					
DB	127 GKFLFIHTHOEMIIGIKKTSGGYRYNDMLAVPIVENTPEKDLKERMARAMTEPDTCAL	186					
DB	: : : : : :~::~ : : : : :~:: : : : :~:: : : : :~:: : : :~:: : : : : :	:					
QY	190 VLVRHGIVYGWGETWEAKTMCECYDYLFDIIVSMKKVLGDPSOLPVGENGVIV	242					
DB	: : : : : :~::~ : : : : :~:: : : : :~:: : : : :~:: : : :~:: : : : : :	:					
DB	187 VLVRHGIVYGWDTWEAKTMCECYDYLFDIIVQMKGHDLPSPVPTKEGIV	239					
DB	: : : : : :~::~ : : : : :~:: : : : :~:: : : : :~:: : : :~:: : : : : :	:					
RESULT 9							
Q6PUX6	PRELIMINARY; PRT: 204 AA.						
ID	O6PJX6						
AC	Q6PJX6;						
DT	05-JUL-2004 (TrEMBLrel. 27, Created)						
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)						
DE	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)						
DE	Hypothetical protein.						
OS	Homo sapiens (Human).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
OX	NCBI_TaxID=9606;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	TISSUE=Placenta;						
RX	MEDLINE=22398257; PubMed=12477932;						
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,						
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,						
RA	Ahtschul S.F., Zeeberg B., Buotow K.H., Schaefer C.F., Bhat N.K.,						
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Heide F.,						
RA	Datchenko L., Marushima K., Farmer A.A., Rubin G.M., Hong L.,						
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,						
RA	Brownstein M.J., Udén T.B., Tohiyuki S., Carninci P., Prange C.,						
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,						
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,						
RA	Richards S., Worley K.C., Haile S., Garcia A.M., Gay L.J., Hulyk S.W.,						
RA	Villalon D.K., Muzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,						
RA	Faney J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,						
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,						
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,						
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,						
RA	Krzywinski M.I., Skaleka U., Smalush D.E., Schnerch A., Schein J.E.,						
RA	Jones S.J., Marra M.A.;						
RT	"Generation and initial analysis of more than 15,000 full-length human						
RT	cDNA sequences.";						
RT	mice mouse cDNA sequences."						

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
DR Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC010133; AAH10133.1; -.
RC Hypothetical protein.
DR InterPro; IPR001303; Aldolase_II_N.
DR Pfam; PF00596; Aldolase_II; 1.
KW Hypothetical protein.
SQ SEQUENCE 204 AA; 22858 MW; 82FC4E97BC6C2853 CRC64;

Query Match 76.0%; Score 1005; DB 2; Length 204;
Best Local Similarity 99.5%; Pred. No. 3.3e-82;
Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 DEIYIAPSGVQKERIQPEDMFVCDINEKDISGSPSKLKKSQCTPLFMNAYTMRGAGAV 113
Db 16 DEIYIAPSGVQKERIQPEDMFVCDINEKDISGSPSKLKKSQCTPLFMNAYTMRGAGAV 75

QY 114 IHTSKAAVMATLLPFGREFKITHQEMIKGIKKTSGGYRYDDMLVVPPIENTPEEKGL 173
Db 76 IHTSKAAVMATLLPFGREFKITHQEMIKGIKKTSGGYRYDDMLVVPPIENTPEEKDL 135

QY 174 KDMAHAMNEYPDSCAVLVRHGVYVWGCTWEKATMCCEYDYLFDIAVSMKKVGLDPSQ 233
Db 136 KDMAHAMNEYPDSCAVLVRHGVYVWGCTWEKATMCCEYDYLFDIAVSMKKVGLDPSQ 195

QY 234 LPVGENGIV 242
Db 196 LPVGENGIV 204

RESULT 10
AAH10133
ID AAH10133 PRELIMINARY; PRT; 204 AA.
AC AAH10133
DT 02-MAR-2004 (TremBLrel. 27, Created)
DT 02-MAR-2004 (TremBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TremBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters K.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010133; AAH10133.1; -.
KW Hypothetical protein.
SQ SEQUENCE 204 AA; 22858 MW; 82FC4E97BC6C2853 CRC64;

Query Match 76.0%; Score 1005; DB 2; Length 204;
Best Local Similarity 99.5%; Pred. No. 3.3e-82;
Matches 188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 54 DEIYIAPSGVQKERIQPEDMFVCDINEKDISGSPSKLKKSQCTPLFMNAYTMRGAGAV 113
Db 16 DEIYIAPSGVQKERIQPEDMFVCDINEKDISGSPSKLKKSQCTPLFMNAYTMRGAGAV 75

QY 114 IHTSKAAVMATLLPFGREFKITHQEMIKGIKKTSGGYRYDDMLVVPPIENTPEEKGL 173
Db 76 IHTSKAAVMATLLPFGREFKITHQEMIKGIKKTSGGYRYDDMLVVPPIENTPEEKDL 135

QY 174 KDMAHAMNEYPDSCAVLVRHGVYVWGCTWEKATMCCEYDYLFDIAVSMKKVGLDPSQ 233
Db 136 KDMAHAMNEYPDSCAVLVRHGVYVWGCTWEKATMCCEYDYLFDIAVSMKKVGLDPSQ 195

QY 234 LPVGENGIV 242
Db 196 LPVGENGIV 204

RESULT 11
Q9HS28
ID Q9HS28 PRELIMINARY; PRT; 153 AA.
AC Q9HS28
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE DJ179L10.2 (Similar to CGI-29 protein) (Fragment).
GN Name=DJ179L10.2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore M.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL138810; CAC12642.1; -.
DR InterPro; IPR001303; Aldolase_II_N.
DR Pfam; PF00596; Aldolase_II; 1.
FT NON TER 153
SQ SEQUENCE 153 AA; 16918 MW; D88A0ADE331E12EA CRC64;

Query Match 62.8%; Score 830; DB 2; Length 153;
Best Local Similarity 99.3%; Pred. No. 1.3e-66;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSGCDACGGDCSRRCGAQDKEHPYLIPELCKQFYHLGWVTGTGGGSLKHGDEIYIAP 60
Db 1 MSGCDAREGDCSRRCGAQDKEHPYLIPELCKQFYHLGWVTGTGGGSLKHGDEIYIAP 60

QY 61 SGVQKERIQPEDMFVCDINEKDISGSPSKLKKSQCTPLFMNAYTMRGAGAVIHTHSA 120
Db 61 SGVQKERIQPEDMFVCDINEKDISGSPSKLKKSQCTPLFMNAYTMRGAGAVIHTHSA 120

QY 121 AVMATLLPFGREFKITHQEMIKGIKKTSGGY 153
Db 121 AVMATLLPFGREFKITHQEMIKGIKKTSGGY 153

RESULT 12
Q9VY93
ID Q9VY93 PRELIMINARY; PRT; 227 AA.
AC Q9VY93
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
```


DE CG11134-PA (R561993p).
GN ORFNames=CG11134;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer J.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Betts P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkac R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Fries E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svrtkac R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svrtkac R.,
RA Patel S., Fries E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RN SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RG FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkeley;
RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fries E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003493; AAF48310.1; -;
DR EMBL; AY071553; AAL49175.1; -;
DR InAct; Q9VY93; -;
DR FlyBase; FBgn0030518; CG11134.
DR InterPro; IPR001303; Aldolase_II_N.
DR Pfam; PF00596; Aldolase_II; 1.
SQ SEQUENCE 227 AA; 26011 MW; 7F2E505906CE155D CRC64;
Query Match 59.8%; Score 790; DB 2; Length 227;
Best Local Similarity 70.3%; Pred. No. 8.3e-63;
Matches 149; Conservative 19; Mismatches 44; Indels 0; Gaps 0;
QY 22 EHPRLIPELCKQFVHLGVTGGTGGISLKHGDEIYIAPSGVKERIOPEDMFVCDINEK 81
DB 12 EHPRLIPSLCRQFHLGVTGGTGGISIKYNDIYIAPSGVKERMPEDLFLVQDITCK 71
QY 82 DISGPSKGLKKSQCTPLFMNAYTMRCAGAVIHTHAKAAVNATLLFPGRFPKITHQEMI 141
DB 72 DLQLPPEIKGLKKSQCTPLFMLAYQHRQAGAVIHTHSHAVNATLLWPGKTRCTHLEMI 131
QY 142 KGIKKCTSGGYRYDDMLVPIENTPEEKLKDRMAHMEYDPSCAVLVRHGVYVWG 201
DB 132 KGVYDEADKRYLRVDEELVPIENTPFERLADSMYAAHMEYPCGSAILVRHGVYVWG 191
QY 202 ETWEKAKTMCYDYLFDIAVSMKKVGLDPSQ 233
DB 192 QNWEKAKTMCYDYLFSIAVEMKKAGIDPEK 223
RESULT 13
Q7PS09
ID Q7PS09 PRELIMINARY; PRT; 277 AA.
AC Q7PS09
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DE ENSANGP0000020764 (Fragment).
GN Name=ENSANGP00000018275;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;

```
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB0100846; EAA06311.2; -.
DR InterPro; IPR001303; Aldolase_II_N.
DR Pfam; PF00596; Aldolase_II; 1.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30799 MW; F7F222B46EAD1F42 CRC64;

Query Match 50.8%; Score 671; DB 2; Length 277;
Best Local Similarity 68.3%; Pred. No. 5.4e-52;
Matches 129; Conservative 19; Mismatches 41; Indels 0; Gaps 0;

QY 22 EHPRLVLPCLCKOFVHLGWVTGGTGGISLKHGDEIYIAPSGVOKERIQPEDMFVCDINEX 81
Dy 89 EHPRLVLPCLCKOFVHLGWVTGGTGGISLKHGDEIYIAPSGVOKERIQPEDMFVCDINEX 148
QY 82 DISGPSPSKKLKKKSQCTPLFMNAYTMRGAGAVIHTHSHKAAVMATLLPFGREFKITHQEMI 141
Dy 149 DLQTPPDYKLTKSQCTPLFMNAYTMRGAGAVIHTHSHKAAVMATLLPFGREFKITHQEMI 208
QY 142 KGIKCTGGYRYDDMLVPIIENTPPEKGLKDRMAHANEYPDSCAVLVRHGVYVWG 201
Dy 209 KGIVDYELNRLMYDEELVPIIENTPPEKGLKDRMAHANEYPDSCAVLVRHGVYVWG 268
QY 202 ETWEKAKTM 210
Dy 269 HNWQAKTM 277

RESULT 14
Q7PG25 PRELIMINARY; PRT; 220 AA.
AC Q7PG25;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000023826 (Fragment).
GN Name=ENSANGG00000018275;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelinae.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB0100846; EAA45094.1; -.
DR InterPro; IPR001303; Aldolase_II_N.
DR Pfam; PF00596; Aldolase_II; 1.
FT NON_TER 1
SQ SEQUENCE 220 AA; 25086 MW; 38AD45AE25F3044 CRC64;

Query Match 46.7%; Score 617; DB 2; Length 220;
Best Local Similarity 65.0%; Pred. No. 3e-47;
Matches 119; Conservative 20; Mismatches 44; Indels 0; Gaps 0;

QY 53 GDEIYIAPSGVOKERIQPEDMFVCDINEXDKISGPSPSKKLKKKSQCTPLFMNAYTMRGAGA 112
Dy 35 GDEIYIAPSGVOKERIQPEDMFVCDINEXDKISGPSPSKKLKKKSQCTPLFMNAYTMRGAGA 94
QY 113 VIHTHSHKAAVMATLLPFGREFKITHQEMIKGIKCTGGYRYDDMLVPIIENTPPEKGL 172
Dy 95 VIHTHSHKAAVMATLLPFGREFKITHQEMIKGIKCTGGYRYDDMLVPIIENTPPEKGL 154
QY 173 LKDRMAHANEYPDSCAVLVRHGVYVWGETWEKAKTMCECYDYLFDIAVSMKKVGLDPS 232
Dy 173 LKDRMAHANEYPDSCAVLVRHGVYVWGETWEKAKTMCECYDYLFDIAVSMKKVGLDPS 232

RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB0100846; EAA06311.2; -.
DR InterPro; IPR001303; Aldolase_II_N.
DR Pfam; PF00596; Aldolase_II; 1.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30799 MW; F7F222B46EAD1F42 CRC64;

Query Match 50.8%; Score 671; DB 2; Length 277;
Best Local Similarity 68.3%; Pred. No. 5.4e-52;
Matches 129; Conservative 19; Mismatches 41; Indels 0; Gaps 0;

QY 22 EHPRLVLPCLCKOFVHLGWVTGGTGGISLKHGDEIYIAPSGVOKERIQPEDMFVCDINEX 81
Dy 89 EHPRLVLPCLCKOFVHLGWVTGGTGGISLKHGDEIYIAPSGVOKERIQPEDMFVCDINEX 148
QY 82 DISGPSPSKKLKKKSQCTPLFMNAYTMRGAGAVIHTHSHKAAVMATLLPFGREFKITHQEMI 141
Dy 149 DLQTPPDYKLTKSQCTPLFMNAYTMRGAGAVIHTHSHKAAVMATLLPFGREFKITHQEMI 208
QY 142 KGIKCTGGYRYDDMLVPIIENTPPEKGLKDRMAHANEYPDSCAVLVRHGVYVWG 201
Dy 209 KGIVDYELNRLMYDEELVPIIENTPPEKGLKDRMAHANEYPDSCAVLVRHGVYVWG 268
QY 202 ETWEKAKTM 210
Dy 269 HNWQAKTM 277

RESULT 15
Q6CBB0 PRELIMINARY; PRT; 238 AA.
AC Q6CBB0;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|P47095|Saccharomyces cerevisiae YJR024c.
GN ORENAMES=VALI0C20427q;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boismare A., Boyer E., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; CR382129; CAG82372.1; -.
SQ SEQUENCE 238 AA; 26590 MW; B63E556136278CE9 CRC64;

Query Match 46.0%; Score 608.5; DB 2; Length 238;
Best Local Similarity 54.5%; Pred. No. 1.9e-46;
Matches 122; Conservative 31; Mismatches 62; Indels 9; Gaps 3;

QY 18 AQDKEHPRLVLPCLCKOFVHLGWVTGGTGGISLKHGDEIYIAPSGVOKERIQPEDMFVCD 77
Dy 10 SSDPKHPANLIVELCKLFYDNNWVITGGISIRREGDTVWLAPSGVOKERIQPTDMFVND 69
QY 78 INEKDISGPSPSKKLKKKSQCTPLFMNAYTMRGAGAVIHTHSHKAAVMATLLPFGREFKITH 137
Dy 70 LKSRDYLRRSPT--FKPSACTPLFLSAYTLRDAGACIHTHSHKAAVMATLLP--DKVFKISN 126
QY 138 QEMIKGKIKCTGGYRYDDMLVPIIENTPPEKGLKDRMAHANEYPDSCAVLVRHGV 197
Dy 127 IEQIKAIQPVVSGYLSFFDTPLEIPIIENTPPEKGLKDRMAHANEYPDSCAVLVRHGV 186
QY 198 YVWGETWEKAKTMCECYDYLFDIAVSMKKVGLDPSQLPVGENGI 241
Dy 187 YVWGETWEKAKTMCECYDYLFDIAVSMKKVGLDPSQLPVGENGI 224

Search completed: February 1, 2005, 14:29:28
Job time : 203 secs
```